С. В. Скуратов

ИСПОЛЬЗОВАНИЕ КЛАВИАТУРНОГО ПОЧЕРКА ДЛЯ АУТЕНТИФИКАЦИИ В КОМПЬЮТЕРНЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Введение

Многие сферы деятельности современного общества зависят от функционирования компьютерных информационных систем (КИС). Последние являются составной частью АС и позволяют решать задачи любой сложности за короткое время. Однако в связи с увеличением объема обрабатываемой информации конфиденциального характера возникла необходимость защиты КИС от угроз, в первую очередь от НСД.

Для защиты КИС принято использовать специализированные системы, называемые системами контроля и управления доступом (СКУД). Для повышения их стойкости в процессе аутентификации используются дополнительные факторы, в том числе биометрические.

В данной работе рассматривается возможность использования клавиатурного почерка в системах многофакторной аутентификации. Представлена информация о разработанной автором архитектуре системы аутентификации, используемых решающих устройствах и результатах тестирования системы.

1. Клавиатурный почерк как поведенческая биометрическая характеристика

Клавиатурный почерк относится к динамическим (поведенческим) биометрическим характеристикам, описывающим подсознательные действия, привычные для пользователя. Он характеризует динамику ввода парольной фразы с помощью клавиатуры. Стандартная клавиатура позволяет измерить следующие временные характеристики: время удержания клавиши нажатой и интервал времени между нажатиями клавиш.

Клавиатурный почерк могут характеризовать и другие параметры, описанные в работе [1]: общее время набора парольной фразы, частота возникновения ошибок при наборе, факт использования дополнительных клавиш (использование числовой клавиатуры), особенности ввода заглавных букв (использование клавиши Shift или Caps Lock) и т. д.

Использование клавиатурного почерка не требует установки специальных аппаратных средств и кадров для установки и поддержки, является прозрачным для конечного пользователя, т. е. не причиняет неудобств пользователю и позволяет проводить скрытую аутентификацию. Клавиатурный почерк также позволяет проводить реаутентификацию для подтверждения личности пользователя перед выполнением критичных операций. Кроме того, клавиатурный почерк обладает всеми преимуществами, присущими биометрическим методам аутентификации и описанными в работе [2].

Согласно работе [3], основные сложности в работе с клавиатурным почерком связаны с большим разнообразием навыков набора у пользователей и влиянием физиологических состояний человека на почерк: сонливости, тревоги, плохого самочувствия и т. п. На ритм ввода могут влиять и другие объективные причины: травма кисти или пальцев руки или устройство ввода нестандартного размера, обладающего другой эргономичностью. Все эти факторы должны быть учтены для обеспечения точной и эффективной работы системы аутентификации.

2. Архитектура системы многофакторной аутентификации

Система многофакторной аутентификации реализуется на базе клиент-серверной архитектуры (см. рис. 1). На стороне клиента имеется Java-приложение (Java-апплет), позволяющее считывать биометрические данные, необходимые для процесса аутентификации или регистрации в системе. Вторым аутентификационным фактором является одноразовый пароль (ОП), генерируемый с

помощью внешнего аппаратного устройства, например токена, мобильного телефона, карманного компьютера и т. п. В случае, если парольная фраза является секретной, описанная схема реализует трехфакторную аутентификацию. Тестовый образец состоит из парольной фразы, динамических характеристик ее ввода и одноразового пароля.

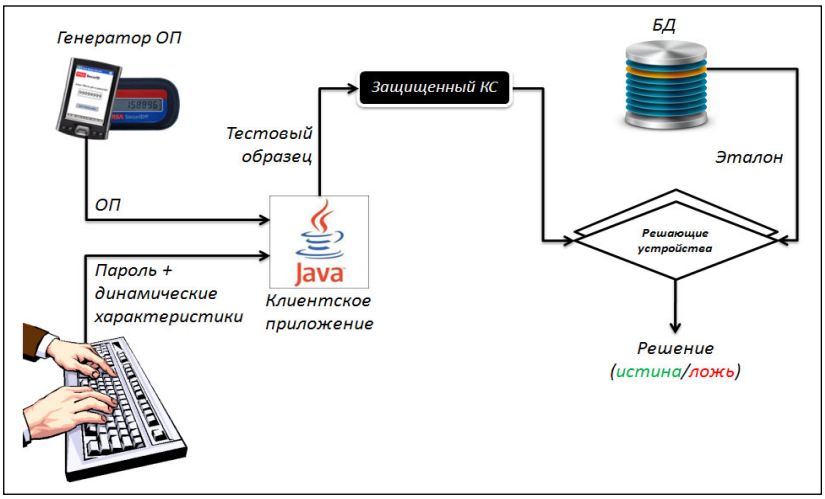

Рис. 1. Архитектура системы

Тестовый образец также содержит имя профиля аутентификации и профиля устройства ввода. Под профилем аутентификации понимается тип набора аутентификационных данных, предоставляемых системе. В наборе может отсутствовать один из факторов, например ОП или динамические характеристики. Это позволяет разграничить доступ к информации с различным уровнем конфиденциальности, т. е. для каждой информационной среды можно использовать свой набор аутентификационных факторов, что одновременно позволяет упростить доступ к менее конфиденциальным ресурсам и использовать полный набор факторов для получения доступа к критически важным конфиденциальным данным.

Под профилем устройства ввода понимается указание на конкретный эталон из БД, используемый в процессе аутентификации. Таким образом, пользователь может пройти регистрацию в системе, используя несколько различных устройств ввода (клавиатур различного типа), например дома и на рабочем месте. Прямое указание на необходимый эталон позволяет существенно повысить точность работы системы.

На стороне сервера находится БД, содержащая эталоны пользовательских аутентификационных данных и векторы инициализации, позволяющие генерировать ОП. На сервере также находятся решающие устройства (РУ), которые реализуют различные алгоритмы сравнения тестовых образцов и эталонов и позволяют обучать эталоны, хранящиеся в БД, с помощью новых тестовых образцов, полученных от аутентифицированного пользователя. Решающие устройства возвращают логическое значение (истина/ложь), являющееся результатом сопоставления предоставленных ему образцов. Несколько РУ с помощью логических операций И/ИЛИ или взвешенных оценок могут быть объединены в сложное решающее устройство.

В целях повышения точности работы системы клиентское приложение на этапе регистрации пользователя осуществляет контроль качества полученных тестовых образцов, не допуская попадания в тестовую выборку некачественных (аномальных) образцов. Кроме того, функционал системы позволяет обучать эталоны тестовыми образцами, с которыми пользователь успешно аутентифицировался в системе, что нивелирует действие физиологического состояния человека на его клавиатурный почерк.

Для защиты от атаки «человека посередине» между сервером и клиентом используется защищенный канал связи (КС), обеспечивающий секретность и целостность передаваемых по нему

данных. ОП главным образом применяется для защиты от атаки повторения, так как динамические характеристики клавиатурного почерка могут быть записаны с помощью вредоносного ПО, например программ-кейлоггеров, а затем повторно воспроизведены.

Данная система многофакторной аутентификации может являться как встраиваемой — в случае если логическое значение, полученное от РУ или их комбинации, используется внутри КИС для нужд ее сервисов, так и автономной, т. е. предоставлять услугу по проверке подлинности пользователя сторонним сервисам. В этом случае требуется организация дополнительного защищенного КС между системой аутентификации и ее клиентами.

3. Решающие устройства

В разработанной системе аутентификации используется несколько типов решающих устройств, работающих на основе меры Хэмминга, метода упорядочивания триграфов, анализа скорости набора парольной фразы и точности ввода. Для всех РУ экспериментальным путем определены их рабочие точки (значения пороговых величин) и значения коэффициентов ложного доступа (КЛД), ложного отказа в доступе (КЛОД) и коэффициента равных ошибок (КРО), соответствующего рабочей точке, в которой КЛД и КЛОД равны. В ходе тестирования было проведено 235200 попыток аутентификации. Исходя из требований к тестовой выборке, изложенных в работе [4], определен ее объем — 25 тестовых образцов для каждого пользователя. Использовались пароли длиной от 20 до 25 символов. Общее число пользователей составило 14 человек.

Мера (дистанция) Хэмминга является мерой различия двоичных векторов одинаковой длины. Обозначим два двоичных вектора b_1 и b_2. Тогда дистанция Хэмминга может быть определена как суммарное число различий между элементами (битами) этих векторов, находящихся на одной и той же позиции. Вычисление меры Хэмминга можно задать и другим образом — это число единиц в векторе $b_1 \oplus b_2$.

Двоичный вектор может быть получен с помощью сравнения тестового образца и эталона. Пусть t — временной вектор из тестового образца и v — временной вектор из эталона. Обозначим h — двоичный вектор, $m(v_i)$ — математическое ожидание величины v_i, $\sigma(v_i)$ – дисперсию величины v_i, $t(N, 1 - P_1)$ — коэффициент Стьюдента, N — число использованных при обучении тестовых образцов, P_1 — заданный (желаемый) КЛОД.

Если, $m(v_i) - t(N, 1 - P_1)\,\sigma(v_i) < t_i < m(v_i) + t(N, 1 - P_1)\,\sigma(v_i)$, (1)

то $h_i = 1$. Иначе $h_i = 0$.

Математическое ожидание величины v_i может быть вычислено по формуле:

$$m_j(v_i) \approx \frac{j-1}{j} m_{j-1}(v_i) + \frac{1}{j} v_{ij}. \quad (2)$$

Дисперсия величины v_i может быть вычислена по формуле:

$$\sigma_j^2(v_i) \approx \frac{j-2}{j-1}\sigma_{j-1}^2(v_i) + \frac{1^i}{j-1}(v_{ij} - m_j(v_i))^2. \quad (3)$$

Мера Хэмминга для векторов h и единичного вектора является мерой различия тестового образца и эталона.

Коэффициент равных ошибок для РУ на основе меры Хэмминга составил 0.015.

Триграфом называются три набранных подряд символа. Длительность триграфа — это время между нажатием первой и третьей клавиш. Например, для пароля «Russian» последовательностью триграфов будет: «Rus» 277; «uss» 255; «ssi» 297; «sia» 326; «ian» 235. Длина триграфов дана в миллисекундах. Триграфы сортируются по возрастанию длительности: «ian» 235; «uss» 255; «Rus» 277; «ssi» 297; «sia» 326. Мера различия между двумя тестовыми образцами определяется как сумма перестановок позиций триграфов в двух отсортированных массивах.

Например, для двух массивов A_1 и A_2, таких что

A_1: «ian» 235; «uss» 255; «Rus» 277; «ssi» 297; «sia» 326;

A_2: «uss» 215; «ian» 258; «Rus» 298; «sia» 306; «ssi» 315.

величина различия будет равна

$$d(A_1, A_2) = \frac{1+1+0+1+1}{mdis} = \frac{1+1+0+1+1}{12} = 0.33. \quad (4)$$

Величина различия нормализована и принимает значения из отрезка [0,1].

Максимальная величина различия mdis между массивами триграфов длины n может быть вычислена по формуле:

$$mdis = \begin{cases} \frac{n^2}{2}, \text{ n – четное} \\ n \cdot \left|\frac{n}{2}\right| + \left|\frac{n}{2}\right|, \text{ n – нечетное} \end{cases}. \quad (5)$$

Для сопоставления тестового образца и эталона вычисляется среднее значение величины различия и сравнивается с пороговой величиной.

Коэффициент равных ошибок для РУ на основе метода упорядочивания триграфов составил 0.11.

Скорость набора парольной фразы (число символов, набираемых в минуту, – SPM) уникальна для каждого пользователя. SPM может быть вычислено по формуле:

$$SPM = \frac{\text{длина парольной фазы} \cdot 60}{\text{общее время набора (в секундах)}}. \quad (6)$$

Для сопоставления тестового образца и эталона вычисляется среднее значение SPM для эталона и с помощью пороговой величины сравнивается со значением для тестового образца.

Коэффициент равных ошибок для РУ на основе анализа скорости набора парольной фразы составил 0.21.

Заключение

Исходя из данных, полученных в ходе тестирования, можно сделать вывод о целесообразности объединения рассмотренных решающих устройств в одно устройство с помощью логических операций И и ИЛИ с использованием подхода, описанного в работе [5]. Учитывая коэффициенты РУ предлагается следующая схема комбинированного решающего устройства:

РУ (Мера Хэмминга) ∧ [РУ (Упорядочивание триграфов) ∨ РУ (Скорость набора)]. (7)

где РУ(<метод работы>) – результат сравнения (истина или ложь) тестового образца и эталона решающим устройством, работающим по методу <метод работы>.

По результатам тестирования коэффициент ложного доступа объединенного РУ составил 0.004, а коэффициент ложного отказа в доступе – 0.064.

Полученные значения ошибок позволяют утверждать, что клавиатурный почерк может использоваться в качестве дополнительного фактора аутентификации в СКУД или средства реаутентификации. Разработанная архитектура многофакторной аутентификации обладает большой гибкостью, что позволяет использовать ее в различных прикладных областях, в том числе в тех, где существует необходимость в контроле доступа к данным, имеющим различную степень конфиденциальности.

СПИСОК ЛИТЕРАТУРЫ:

1. *Ilonen J.* Keystroke Dynamics // Lappeenranta University of Technology. 2008.
2. *Checco J. C.* Keystroke Dynamics and Corporate Security // WSTA Ticker Magazine. 2003.
3. *El-Hadidi Kamal M.* Biometrics. What and How. 2007. URL: http://www.net-security.org/dl/articles/Biometrics.pdf.
4. *Иванов А. И.* Нейросетевые алгоритмы биометрической идентификации личности. М.: Радиотехника, 2004.
5. *Bolle M. B., Connell J. H., Pankanti S., Ratha N. K., Senior A. W.* Guide to Biometrics. NY., 2004.